From logic to language

Basile, Valerio

STELLINGEN

behorende bij het proefschrift

From Logic to Language
Natural Language Generation from Logical Forms

van

Valerio Basile

1) Logical Forms are a good candidate for the role of input for Natural Language Generation, both for being a theoretically sound formalism, and for the practical implications.

2) Discourse Representation Structures give the possibility to produce a surface form for each discourse referent separately, as well as the surface form for an entire sentence or even discourse, by virtue of composition.

3) Discourse Representation Graphs retain the expressive power of Discourse Representation Structures, while simplifying the word-level alignment.

4) The Groningen Meaning Bank provides a large enough quantity of text aligned with logical forms to be used as training material for supervised stochastic methods for NLG. The crowdsourcing approach to annotation ensures that its quality constantly improves over time.

5) WordNet is a valuable resource for generating expressions of natural language, despite not being a full-fledged ontology, because it provides sets of lemmas for each concept to generate, thus transforming the problem of lexicalization into the simpler problem of lexical choice.

6) The lexicalization of the concepts of a Discourse Representation Graph is choosing a lemma and generating the morphology.

7) The task of predicting whether or not a part of a logical form participates in the generation process is easy, to order them correctly is hard, especially with respect to event referents.

8) The evaluation of the output of a natural language generation pipeline based on string matching is too strict, counting as errors productions that are arguably acceptable. This applies to the evaluation of the word order as well as the choice of lexemes.

9) Academic code is like a *frikandel*: monolithic, mysterious, and hard to process.

10) "Gezellig" is een gezellig woord.